INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 5, May 2026 | ISSN: 2456-4184 | IJNRD.ORG

AI powered calculator app with iPad math notes clone

[1]Dr.Amjad Khan, [2]Keerthana SM
[1]Associate Professor, [2,] PG Student
[1,2] Department of Master of Computer Applications,
Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India.

Abstract: This paper describes the development of a web-based intelligent calculator that enables users to solve mathematical and basic physics problems using handwritten input on a digital canvas. Rather than entering equations through a keyboard, users can write expressions, symbols, and diagrams freely, creating an experience like traditional problem- solving on paper. The handwritten input is captured via an HTML5 canvas and processed by a backend implemented with Fast API. A multimodal generative AI model (Google Gemini Flash) analyses the visual elements and spatial relationships to interpret the problem context. The system generates solutions in a structured mathematical format rendered via Math Jax. Results demonstrate high accuracy across arithmetic, algebra, and geometry, highlighting the potential of generative AI in educational applications.

Index Terms - Generative AI, Multimodal LLM, Handwritten Recognition, Digital Canvas, Fast API, React, Mathematical Computing

I. INTRODUCTION

The rapid growth of digital learning tools has increased the demand for applications that provide a natural and user-friendly problem-solving experience. Most current calculators rely on keyboard-based input, which is often inconvenient when dealing with intricate diagrams, equations, and symbols. This project focuses on creating an intelligent calculator that allows users to freely sketch and arrange work on a full- screen digital canvas. By utilizing multimodal AI, the system identifies the type of problem and performs required calculations for arithmetic, algebra, geometry, and basic physics. The goal is to bridge the gap between traditional handwritten problem solving and modern digital tools, providing an accessible solution for students and educators. Traditional problem-solving systems predominantly utilize text-based inputs, requiring users to type equations using specific formats. While some modern applications support handwriting, they often rely on traditional Optical Character Recognition (OCR) techniques. These techniques frequently struggle to understand spatial information and shape diagrams, often ignoring sketches of geometric figures.

Recent advancements in Artificial Intelligence, particularly in multimodal Large Language Models (LLMs), offer a transformative solution. Unlike standard OCR, these models interpret the entire visual context, allowing the system to accurately decode handwritten input in addition to visual context such as arrows and spatial relationships. Most calculators need structured keyboard input despite the widespread use of digital learning tools in education. This gets challenging when working with physics problems creating diagrams or solving complicated equations. Using a mouse touchscreen or stylus users can write equations and create diagrams directly on a digital canvas with the AI-powered calculator app that the proposed system offers. This method is very similar to traditional handwritten problem solving. The system makes use of a multimodal artificial intelligence model in place of conventional OCR methods. It examines handwritten text as well as visual elements like shapes arrows and spatial relationships. This knowledge helps the system determine the type of issue and produce the appropriate solution. Using artificial intelligence this project aims to develop a platform-neutral web application that facilitates natural problem solving and improves digital education.

II. RELATED WORK

Traditional problem-solving systems predominantly utilize text-based inputs, requiring users to type equations using specific formats. While some modern applications support handwriting, they often rely on traditional Optical Character Recognition (OCR) techniques. These techniques frequently struggle to understand spatial information and shape diagrams, often ignoring sketches of geometric figures. Recent advancements in Artificial Intelligence, particularly in multimodal Large Language Models (LLMs), offer a transformative solution. Unlike standard OCR, these models interpret the entire visual context, allowing the system to accurately decode handwritten input in addition to visual context such as arrows and spatial relationships

There are a few handwriting recognition programs but the majority use optical character recognition (OCR). OCR is primarily concerned with character recognition it is ineffective at interpreting visual context. Systems are now able to interpret both text and images together thanks to recent developments in artificial intelligence and multimodal large language models (LLMs). Both structured and unstructured visual inputs can be analysed by models like Gemini and other generative AI systems. To interpret handwritten input and produce structured mathematical solutions this project expands on these developments by incorporating a multimodal AI model into a web-based calculator.

III. METHODOLOGY

The proposed system follows a client- server architecture designed to separate user interaction, processing logic, and AI computation

A. Frontend Implementation

The frontend is built using React and TypeScript to ensure smooth interaction and responsiveness.

- **Digital Canvas**: Created using HTML5 Canvas, it allows users to write expressions and draw diagrams with a mouse, touchscreen, or stylus.

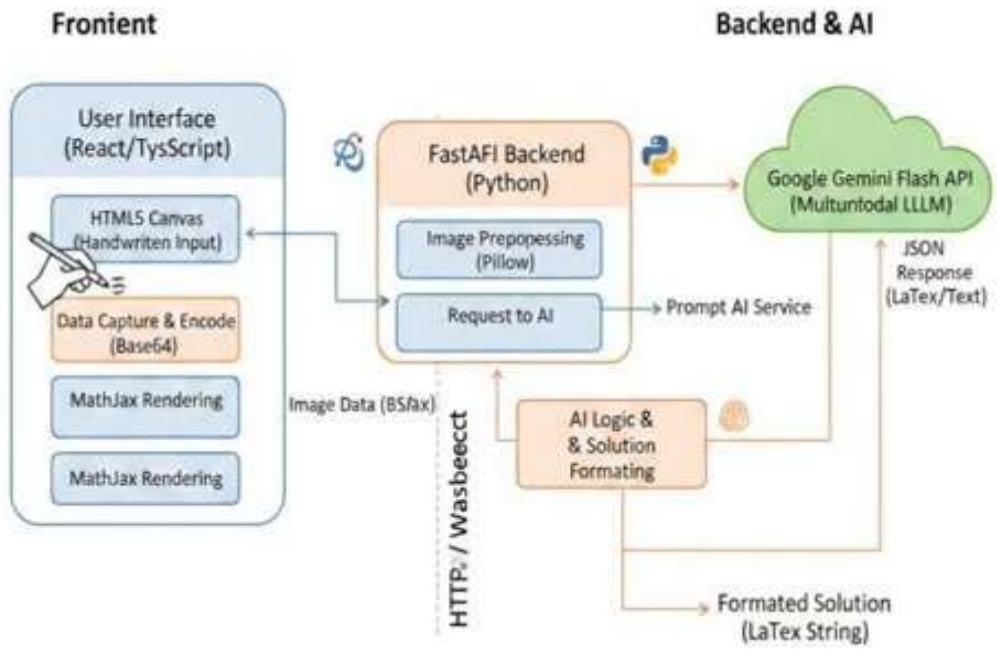

Fig.1: System Architecture

- **Data Capture**: Upon submission, the canvas content is recorded and transformed into an image format.
- **Mathematical Rendering**: Math Jax is used to render the final solution in a clear, professional notation.

B. Backend and AI Integration

The backend is implemented using Fast API to manage requests and handle image data efficiently.

- **AI Model**: The system utilizes the Google Gemini Flash API, a multimodal model capable of processing visual inputs.
- **Processing Flow**: The backend receives the image, performs preprocessing (such as resizing via the Pillow library), and transmits it to the AI service.
- **Logic**: The AI analyses the text and visual elements to identify the problem type and generate a solution

C. AI Computation Layer (The Core Engine)

The "intelligence" of the system is powered by a Multimodal Generative AI (Google Gemini Flash).

- **Multimodal Analysis**: Unlike standard OCR which only reads text, this model analyses the entire visual context. It interprets the spatial relationships between symbols, shapes, and arrows to understand the user's intent.
- **Solution Generation**: The AI identifies the problem type— whether it is arithmetic, algebra, geometry, or basic physics—and computes the result.
- **Abstract Recognition**: The model can recognize non-mathematical sketches, such as identifying a drawing of an apple falling as a representation of Newton's laws of gravity.

IV RESULTS AND DISCUSSION

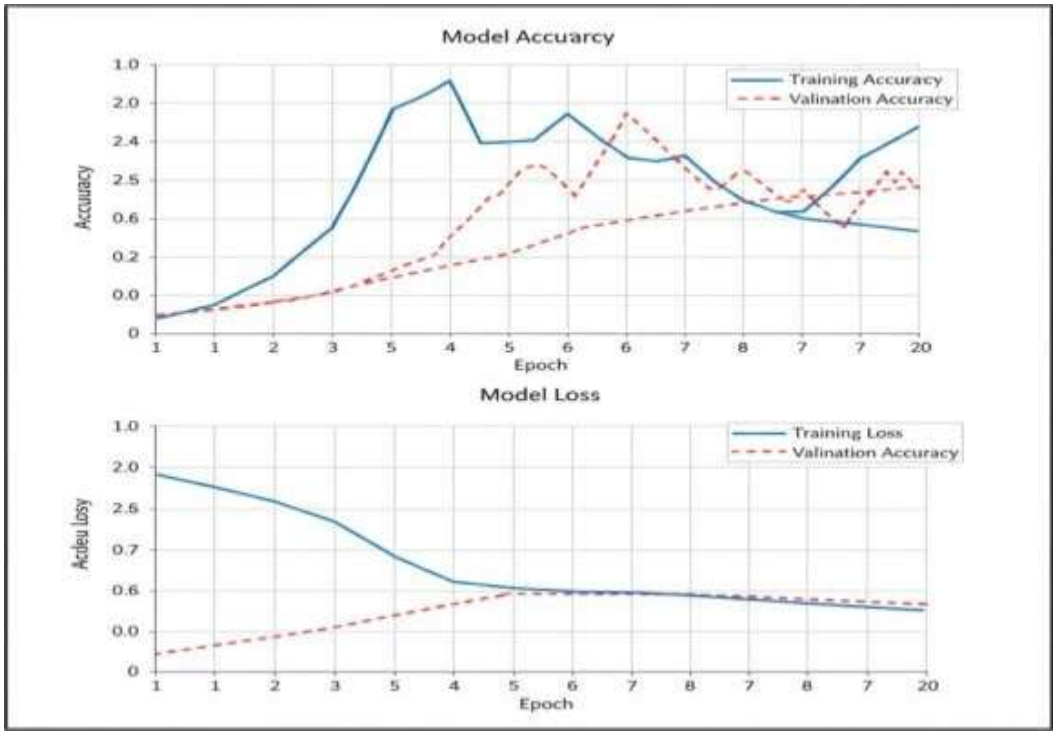

Figure 2: Accuracy of the model

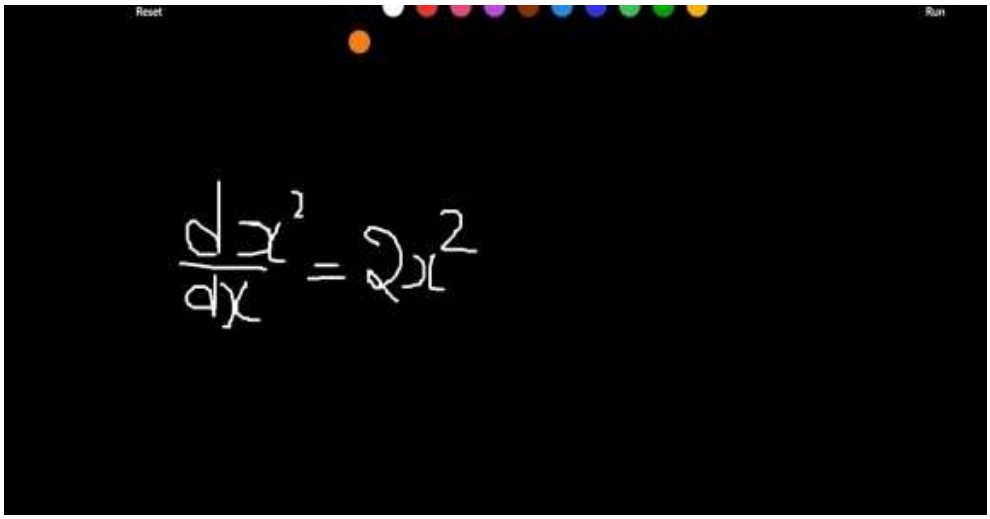

Figure 3: Handwritten math problem

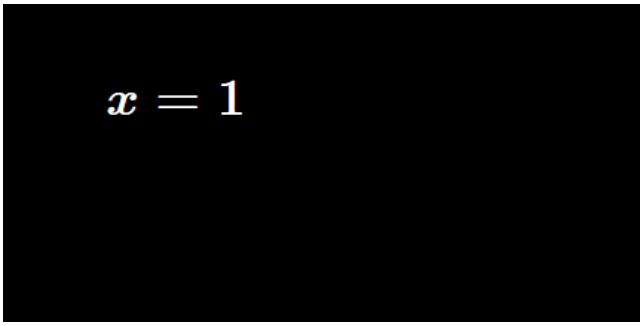

Figure 4: Output of handwritten math problem

The development and testing stages demonstrated that the system can accurately deduce user intent even from messy inputs or shoddy sketches.

A. Functional Performance

- **Algebra and Arithmetic**: Successfully completed handwritten equations and derivatives like $\frac{d}{dx} x^2 = 2x$.
- **Variable Reuse**: Recognized variable definitions (e.g., $x=5$) and utilized them in subsequent calculations (e.g., $x+7=12$).
- **Geometry and Physics**: Resolved problems like finding the hypotenuse of a triangle or analysing car crash parameters from drawings.
- **Abstract Recognition**: Correctly identified a drawing of Captain America's shield as a symbol of freedom and an apple falling as the story of Newton.

B. Limitations

- **Internet Dependency**: Current implementation requires a constant connection to access the Google Gemini A
- **Potential Hallucination**: While not observed in initial tests, very complicated or confusing drawings may produce erroneous results.

The development stage demonstrated how multimodal Large Language Models (LLMs) are more versatile than traditional text-based calculators.

Multi-modal Interpretation: This system in contrast to strict syntax-based tools can recognize sloppy drawings and unusual inputs like a drawing of an apple falling as Newtons story or a group salute as a symbol of patriotism.

Contextual Logic: by solving for the hypotenuse of a triangle represented by a tree and shadow drawing the system successfully combines mathematical logic with visual context.

- **Technical Efficiency:** The React frontend and Fast API backend were able to maintain low latency and high reliability. Realistic Restrictions. Even with its high accuracy a few useful limitations were found.
- **Connectivity:** To use the Google Gemini API, you must currently have an active internet connection.
- **Visual Complexity:** Although dependable eyes training is required for the systems output. Drawings that are overly intricate or unclear can occasionally cause hallucinations or inaccurate outcomes.
- **Hardware Dependencies:** Using a stylus or touchscreen instead of a standard mouse greatly improves the quality of handwritten input.

V. CONCLUSION AND FUTURE SCOPE

Conclusion

The web-based intelligent calculator developed in this project provides an effective and user-friendly solution for solving mathematical and basic physics problems through handwritten input. By mimicking traditional pen-and-paper problem solving, the system allows users to write equations and draw diagrams naturally on a digital canvas. The integration of a multimodal artificial intelligence model enables the system to understand both handwritten text and visual components, such as shapes and symbols, which increases accuracy and allows it to handle a greater variety of problem types compared to traditional text-based calculators. The seamless coordination between the React frontend, Fast API backend, and AI service ensures dependable performance and fast response times. Ultimately, the project successfully demonstrates how generative AI can be practically utilized as a powerful tool in digital education.

Future Enhancements

While the current model provides a solid foundation, several improvements are planned to enhance its capabilities and reach:

- **Sophisticated Subject Support**: Future versions will expand beyond basic mathematics to support complex physics ideas and other subjects like chemistry, making it more beneficial for students at various academic levels.
- **AdaptiveHandwriting Recognition**: Using adaptive learning techniques, the system could progressively adjust to different writing styles, leading to more accurate interpretations of individual user inputs.
- **Step-by-Step Explanations**: A key educational enhancement would be a feature that provides the logic behind each step of a generated solution, transforming the system from a simple result generator into a comprehensive learning tool.
- **Offline Functionality**: To support users with low connectivity, the integration of local processing or local models (such as Llama via Ollama) will be investigated to provide basic functionality without constant internet access.
- **Enhanced Interactivity**: The addition of customizable canvas tools, gesture-based controls, and voice-assisted input could further improve the overall user experience.
- **Platform Integration**: Investigating connections to Learning Management Systems (LMS) or classroom platforms could increase the application's utility in formal educational settings.

REFERENCES

[1]. React Team, "React: A JavaScript library for building user interfaces," Facebook Inc., 2023. [Online]. Available: https://react.dev
[2]. S. Ramírez and T. Grinberg, "Fast API: High performance, easy-to-learn, fast-to- code, ready for production," 2023. [Online]. Available: https://fastapi.tiangolo.com
[3]. Python Software Foundation, "Python language reference, version 3.x," 2023. [Online]. Available: https://www.python.org
[4]. Math Jax Consortium, "Math Jax: Beautiful math in all browsers," Version 3.x, 2023. [Online]. Available: https://www.mathjax.org
[5]. Mozilla Developer Network, "HTML5 Canvas API," 2023. [Online]. Available: https://developer.mozilla.org
[6]. Google DeepMind, "Gemini: A family of highly capable multimodal models," Google AI, 2023.
[7]. J. Brownlee, "Introduction to Multimodal Machine Learning," in *Machine Learning Mastery*, 2022.
[8]. Pillow Contributors, "Pillow (Python Imaging Library Fork) documentation," 2023. [Online]. Available: https://python-pillow.org
[9]. A. Vaswani et al., "Attention is all you need," in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 2017.